
心智与计算, Vol.1, No.2 (2007), 281-290

文章编号: MC - 2007-025

收稿日期: 2007-05-30

出版日期: 2007-06-30

© 2007 MC - 厦门大学信息与技术学院

古琴减字谱图像的文本切分与提取

潘知梹, 周昌乐*

(厦门大学智能科学与技术系, 福建 厦门 361005)

dozero@xmu.edu.cn

摘要: 古琴打谱是一项专业性很强又费时费力的工作, 如何利用先进的人工智能技术来对古琴谱中的谱字进行自动识别解读, 哪怕是辅助性的, 对于古琴打谱事业的发展无疑会起到重要的促进作用, 从可以间接地为保护与弘扬古琴文化作贡献。本文通过研究古琴减字谱这种特殊文本的特点, 提出了不同于普通 OCR 光学识别软件的文本切分方法, 内容涉及纸质古琴谱原始扫描图的图像预处理, 古琴混合谱中简字谱的行切分与提取, 单行简字谱中的谱字切分与提取等算法及其实现。这样就为实现古琴谱进一步的释读提供一种前期处理方法。由于古琴谱字的特殊性, 以及所提出方法的通用性, 该方法对于丰富汉字文本的切分方法也有着一定的学术意义。

关键词: 古琴打谱; 减字谱; 图像处理; 文本切分

中图分类号: TP 311.5

文献标识码: A

Text Segmentation and Extraction from Images of Guqin Jianzi Pu

PAN Zhi-xiao, ZHOU Chang-le

(Department of Intelligence Science and Technology, Xiamen University, Xiamen 361005, China)

dozero@xmu.edu.cn

Abstract: To transcribe Guqin Jianzi pu (reduced notation) is a highly-professional, time-consuming and laborious work. If to identify and transcribe characters in Jianzi pu can be pursued automatically by means of advanced artificial intelligence techniques, even if the work is subsidiary it will play an important role in the development of Guqin Jianzi pu transcription and contribute to protect and promote Guqin culture as well. This paper investigates the attributes of Guqin Jianzi pu and proposes an OCR method for segmentation and extraction notations from images of Jianzi pu, including algorithms and implementations of pretreatment of initial scanned images of Jianzi pu, segmentation and extraction of Jianzi pu from mixed notation of Jianzi pu and stuff notation and segmentation and extraction of characters from a single line Jianzi pu. This work orients

*通讯作者

to the particularities of Guqin Jianzi pu and performs universally, so to some extent it also enriches the research on Chinese character segmentation.

Key words: Transcription of Guqin Jianzipu; Jianzipu; Image processing; Text Segmentation

1 引言

古琴，又称“文武七弦琴”，是中华民族最早的弹弦乐器，有三千多年的历史，是中华优秀传统文化之瑰宝。联合国教科文组织 2003 年 11 月 7 日在其巴黎总部宣布了世界第二批“人类口头和非物质文化遗产代表作”，中国的古琴艺术名列其中。

古琴是保留中国古代音乐风貌最多的乐器。历代传下来的古琴谱集约有 150 多种，琴曲 600 多首，而同一琴曲不同版本的乐曲总数则达 3000 多首。但现在一般琴人所广为弹奏的琴曲，亦不过 100 余首而已。

古琴音乐保存了很多古代的遗声。琴曲的打谱不仅将向人们展示出琳琅满目、丰富多彩的古代琴乐作品，为继承发展传统音乐提供了生动的实例，并且对于更深入地研究中国古代音乐史以及古代调性、调式、音阶、律制的生成、变化与发展，有着十分重要的意义。因此，古琴曲的打谱，是古琴音乐遗产研究中的一项重要内容，对于研究中国传统音乐与文化，具有很高的学术价值。但是，古琴的生存状况不容乐观。懂古琴的人本来就很少，加上古琴很少在公众场合演奏，致使大众对古琴的熟悉和重视程度很低。

为了弘扬古琴艺术，厦门大学艺术认知实验室的古琴研究小组希望在收集琴谱的基础上，通过人工智能技术来自动提取琴谱中的减字谱字加以识别，然后通过减字编码输入和古琴音色合成技术，实现一定程度古琴打谱的自动进行目标。

本文所开展的研究就是古琴减字谱中的减字切分与提取部分，旨在为琴谱的自动化输入提供一种方便的手段。显然这种将人工智能技术应用于古琴谱的解读，从而使古琴谱研究者可以从繁琐的事务劳动中解脱出来，对于保护与传承古琴文化，是有着积极意义的。这也就是我们从事古琴计算化研究的初衷与价值所在。

目前国内外尚没有专门从事古琴计算化研究的专业实验室，也尚无开展减字谱切分识别方面研究工作。因此，我们的工作对于填补古琴谱自动识别的空白，也是具有深远的学术意义和应用价值的。

2 减字谱及其文本切分策略

古琴减字谱是一种特殊的、为传统琴曲所专用的记谱方法，至少已有 1500 多年的历史。唐代以前用文字记述弹琴的指位和弦位，称文字谱。唐代琴人曹柔在文字谱的基础上创建了减字法，将原来的文字谱减化、缩写，并形成了较为系统的古琴减字谱，并一直为明清两代沿用。减字谱由汉字的部首、数字和一些减笔字拼合而成。减字谱中的一个谱字通常可分为上下两大部分：上半表示左手指法及位置，下半表示弦次及右手指法。减字谱主要是记录弦位和徽位、左右手的弹奏方法，但不直接记录音高。用

减字谱记录而传承至今的古琴谱有 150 多种。值得注意的是,减字谱可以反映三种律制,有的方面用五线谱则无法体现。因此,减字谱记录古琴音乐的完整性和科学性,是西方五线谱等记谱方法所无法取代的。这也就是古琴减字谱的魅力所在。可以说减字谱不仅仅是古琴曲的一种记谱方法,也是代表了一种特殊的音乐思维方式,更是研习古琴文化、传承古琴音乐的一种不可或缺的手段。



图 1 减字谱汉字示例

Fig.1 Jianzipu demonstrations of Chinese characters

由于古琴曲谱是以减字谱的形式保存下来,而减字谱采用的又不是常规汉字,而是复合而成的不规则字形,这就使得目前已有的各种 OCR 光学文字识别软件都无法对其进行准确的切分和识别。为了解决减字谱谱字的释读,首先需要解决减字谱文本的切分问题,给出一个个具体谱字有效的切分方法。

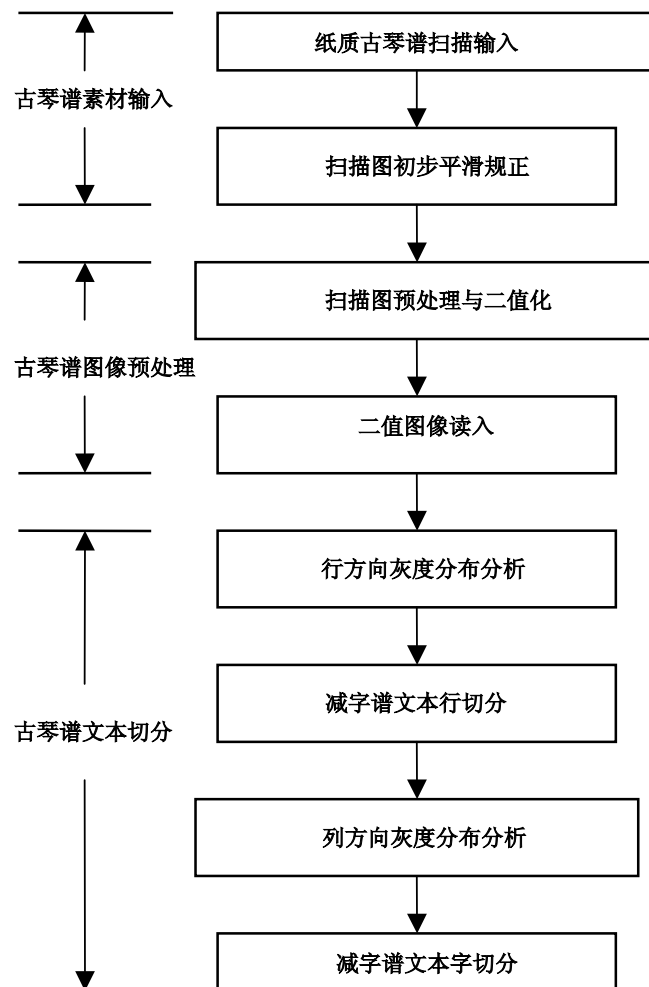


图 2 古琴谱文本切分研究策略流程

Fig.2 Text segmentation strategy process of Guqin Pu

图 2 是我们提出的一种古琴减字谱文本切分的策略流程,主要是通过应用图像阈值处理技术和文本

行字切分技术,实现古琴谱中的五线谱与减字谱的分割,并最终切分出一个个具体的二值表示的减字谱字图像,作为单字识别的输入数据。

图中的设计流程大致是:纸质古琴谱→(图像输入设备接收)→原始扫描图→(图像处理技术:如滤波和变换技术的平滑规正)→保真度好的数字灰度图像→(二值化)→黑白图(也即数字二值图像)→(行切分)→与五线谱分离开的减字谱→(字切分)→单个谱字。实现该流程各个主要环节的效果示意图见图3。



图3 项目实施功能流程示意

Fig.3 The sketch map of projects implementation and function processes

在上述流程中,主要研究步骤包括三个部分:(1)纸质古琴谱原始扫描图的图像预处理,主要包括平滑规正和二值化处理。通过分析古琴谱黑白图像的灰度分布特点,采用统计灰度直方图的阈值化技术,将初始图像转化为单纯的二值图像,增强有关信息的可检测性和最大限度地简化数据,从而改进检测和识别的可靠性,为进一步的图像处理提供方便。(2)古琴混合谱的简字谱行切分与提取,采取的方法是通过程序实现行灰度投影函数,并通过研究其特点,采用统计行段和间段长度的方法,进行字行与空白行的切割。(3)单行简字谱中的谱字切分与提取,采取的方法是通过程序实现字灰度投影函数,采用类似行切分的算法,统计字符宽度与间隔宽度,进行字与字间的定位与切分。

3 古琴谱图像预处理与二值化

简单把书写的汉字形体看作是由黑色的字形和白色的背景,不但不会影响对汉字的识别,而且还会大量的提高时间效率^[1]。因此,作为手写汉字切分的第一步,就是要将获得的数字灰度图像转变为数字二值图像。这个过程为二值化过程,采用的技术通常称为阈值化技术。在众多的阈值化技术中,根据古琴谱图像的特征,我选用较为普遍的灰度直方图法对其进行二值化^[2]。

灰度直方图(histogram)是灰度级的函数,描述的是图像中每种灰度级像素的个数,反映图像中每种灰度出现的频率。横坐标是灰度级,纵坐标是灰度级出现的频率。灰度直方图用于显示图像的灰度值分布情况,表明在每一灰度级有多少个像素,是数字图像处理中最简单和最实用的工具。

灰度级直方图是阈值面积函数的导数的负值,因此若图像具有L(通常L=256,即8位灰度级)级灰度,则大小为M×N的灰度图像f(x,y)的灰度直方图hist[0…L-1]可用如下计算获得:

- (1) 初始化 hist[k]=0; k=0, …, L-1
- (2) 统计 hist[f(x,y)]++; x=0, …, M-1, y =0, …, N-1

(3) 归一化 $\text{hist}[f(x,y)]/M*N$

图 4 给出了一个古琴谱灰度直方图计算的示例^[3]。通过对数字灰度图像进行灰度直方图统计, 然后根据灰度分布具体情况, 可以动态选择阈值来进行图像的二值变换。一般取灰度直方图中两峰值之间的谷值作为阈值。在汉字处理中, 所选择的合适的阈值, 将使得大于该值的图像部分形成汉字的字型, 而小于该值的图像部分构成汉字背景^[4]。

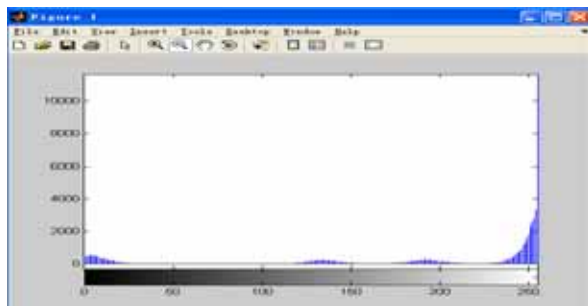


图 4 古琴谱灰度直方图示例

Fig.4 Grey-level Histogram demonstration of Guqin Pu



图 5 (a) 二值化前的古琴乐谱灰度图像片段

Fig.5(a) Fragment of grey-level image of
Guqin Pu before binary



图 5 (b) 选取阈值为 225 时得到的二值化图像

Fig.5(b) Binary images with threshold 255



图 5 (c) 选取阈值为 50 时得到的二值化图像

Fig.5(c) Binary images with threshold 50



图 5 (d) 选取阈值为 150 时得到的二值化图像

Fig.5(d) Binary images with threshold 150

图 5 给出的是经过不同阈值二值化处理的效果对比。从图中可知, 对图 5 (a) 所示的初始灰度图像进行二值化处理, 当选取阈值为 225 时, 得到的图像过于嘈杂; 而选取阈值为 50, 得到的图则略为浅淡; 仅当取阈值为 150 时, 得到的二值图像是最清晰分明的, 故采用阈值 150 对琴谱图进行二值化处理, 使其转化为如图 5 (d) 所示的完全黑白图像。

4 行切分方法与算法^[5]

图像预处理后, 获得的文本二值图像不但包含了组成减字谱的一行行汉字, 也包含了五线谱。这就

需要采用一定的处理技术,将五线谱行与减字谱行切分开来。

从根本上讲,行切分本身就已利用了文本整体局势的信息,所以必须从整体文本的行分布角度出发实现行切分。不失一般性,可以设文本的二值图像为 $f(i, j)$, 图像大小为 $M_x \times M_y$ 。定义 $f(i, j)$ 在 i 行上的投影函数为 $g(i) = \sum_{j=1}^{M_y} f(i, j)$, $i=1, 2, \dots, M_x$ 。 $g(i)$ 如图 6 所示。



图 6 古琴谱图像行灰度投影示例

Fig.6 Grey-level demonstration of Image bank of Guqin Pu

$g(i)$ 反映了文本图像按行灰度累积密度分布情况。通过分析 $g(i)$ 的分布规律,可以获得文本图像按行书写的排列情况,从而确定每行汉字的首行和尾行下标。

通常,如果 i 行图像处于行间隔,则 $g(i)$ 几乎为 0,则 $g(i)$ 至少大于一个字的平均投影密度。这样,可以在单字平均投影密度值和 0 间选择一个阈值对 $g(i)$ 进行二值处理,形成二值序列 $g_1 g_2 \dots g_{M_x}$, 其中 $g(i)$ 为 0 或 1。序列中连续为 1 的子列为行段,连续为 0 的子列为间段,行段或间段中 1 或 0 的个数为段长。于是,古琴谱的行切分可按如下算法来实现:

- (1) 生成文本图像的二值序列 $g_1 g_2 \dots g_{M_x}$, 并求出全部行段和间段;
- (2) 分别求出最大行段长度和平均长度, 如果最大行段长度远大于汉字大小, 进行文本图像的旋转正规化处理, 转(1), 否则
- (3) 令 $k=1, i=1, l=1$;
- (4) 对段 k 进行处理:
 - a. 若段 k 为间段, 令 $i=i+\text{段长 } k$;
 - b. 若段 k 为行段且段长 $k > 2/3$ 行段均长, 则令分行 $l = \langle i, i+\text{段长 } k \rangle$ 及 $i=i+\text{段长}$;
 - c. 若段 k 为行段且段长 $k \leq 2/3$ 行段均长, 则确定最大行段下标 $k' > k$, 满足段长 $j \leq \text{行段最大长度}$. 并令分行 $l = \langle i, i+\sum \text{段长 } j (j=k \wedge k') \rangle$, $i=i+\sum \text{段长 } j, (j=k \wedge k')$, $k=k'$;
 - d. $l=l+1, k=k+1$;
- (5) k 小于段总数, 转(4);
- (6) 形成分行 $j = \langle \text{分行首位 } j, \text{分行末位 } j \rangle, j=1, 2, \dots$ 。

在图像处理的具体程序实现中,是把图像当作是像素点阵来处理。每个图像都是一个矩阵。图像的灰度累积可以通过循环语句统计矩阵中每个点的数值来实现。再通过统计零值与非零值,可以区分出行段与间段,并以向量存储段的起始与终止位置,最后加以切割。

5 字切分方法与算法^[5]

设字行图像为 $L(i, j)$, $i=1, 2, \dots, w$; $j=1, 2, \dots, My$, 其中 w 为该字行的行宽, 可令 L 的字列投影函数为 $h(j)=\sum L(i, j)$, $i=1 \sim w$. $h(j)$ 如图 7 所示。



图 7 古琴谱图像列灰度投影示例

Fig.7 Grey-level demonstration of Image out of Guqin Pu

通过在 0 和单笔画平均厚度之间取阈值将 $h(j)$ 进行二值变换, 可形成二值序列 $h1h2 \dots hMy$. 同样可以称连续为 1 的子序列为字段, 而连续为 0 的子序列为间段, 并定义连续 1 或 0 的个数为段长。

根据古琴简字谱的特性可知, 其字与字之间的间隔较大, 不存在汉字的粘连, 故不需进行粘连检测和左右部首的合并搭配。可认为每个字段至多只含有一个汉字。减字谱图像的行投影比较整齐, 除了投影的波峰处对应文本行外, 在每一行和下一行之间有较宽的一段投影值为 0, 对应了空白处, 给行切分留下了较大的余地。于是, 古琴谱字行的字切分可按如下算法来实现:

(1) 根据 $L(i, j)$ 求出二值序列 $h1h2 \dots hMy$, 并规定第 k 个段左边的间段为第 k 个间段;

(2) 求出平均间段长度和长度阈值为 $1/3(w + \text{最大字段长度} + \text{平均字段长度})$;

(3) 令 $k=1$, 间段长度 $l = \text{间段均值}$, $j = \text{字段 } 1 \text{ 起始位置}$, $l=1$;

(4) 寻找使下式成立的最小的 $k' \geq k$;

$\sum (1/2(\text{间段长度 } i + \text{间段长}(i+1)) + \text{字段长度 } i) > \text{长度阈值}$, $i=k \sim k'$;

(5) 计算 $m = \text{字段长度 } k + \sum (\text{间段长度 } i + \text{字段长度 } i)$, $j=k+1 \sim k'$;

(6) 形成分字 $l = \langle j, j+m \rangle$;

(7) $k=k' + 1$, $j=m+j + \text{间段长度 } k' + 1$;

(8) 如果 $k < \text{字段总数}$, 则 $L=L+1$ 转 (4), 否则

(9) 得到 l 个分字 $i = \langle \text{分字首位 } i, \text{分字末位 } i \rangle$ 。

通过字切分算法, 我们可以获得一个个切分好的单独的谱字图像, 为下一步谱字的识别解读提供规范的输入。

6 古琴减字谱文本切分程序的设计及其实现^[6-8]

整个减字谱文本切分程序系统旨在提取出现有的古琴曲库中曲谱的简字谱, 形成完整的简字序列, 供后期的自动识别使用。图 8 给出了该系统主要的功能模块。

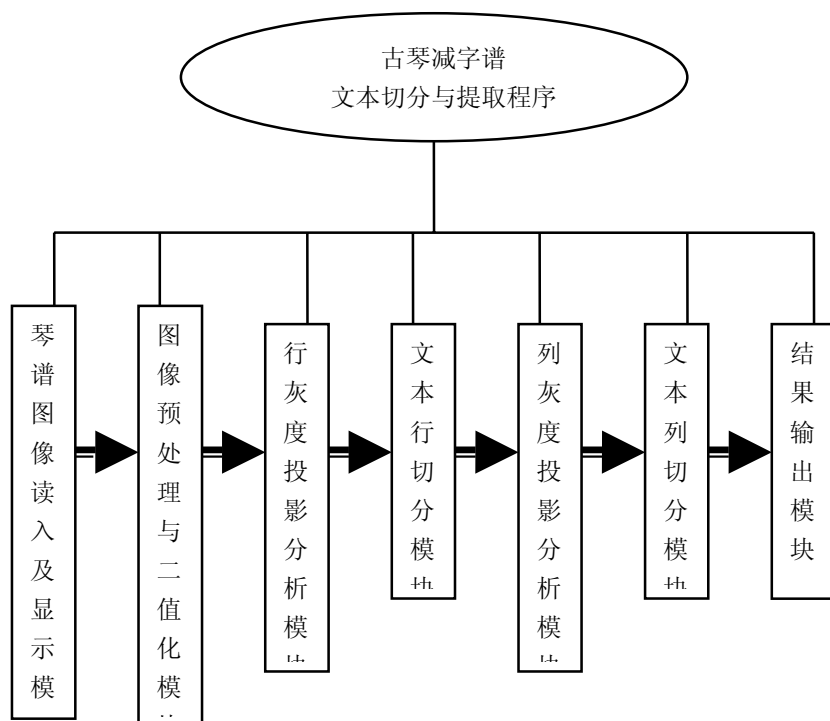


图 8 主要功能模块关系组织图

Fig.8 The organization chart of the main function module relationships

系统采用 MATLAB 6 在 PC 机环境下实现的^[3]。就古琴减字谱识别切分而言，我将系统与现有主要的 OCR 软件做了比对实验，有如下结果。

采用汉王 OCR 5.0 增强版来处理，实验结果发现只能处理初步的行分析，而不能进行字分析。并且分析结果中出现大量莫名的线段标识，比较混乱不清。总之其没有行切分与字切分功能。参见图 9。

采用尚书七号 OCR 进行实验，结果参见图 10，图中绿色框内为该软件对图片行分析的结果，红色框内为字分析结果。结果可以看到存在把五线谱行和字行合为一行的现象，即根本未实现正确的行切分。究其原因，是其无法严谨的对连接较密的谱行与字行进行分离识别。更存在把多个减字谱连成一体误分析为单个汉字的情况。我认为可能是由于其算法中默认把间隔小于最大字间距离的几个汉字都当作同个汉字的左右偏旁来看待，导致在应用于古琴谱这种特殊情况时字切分失败。



图 9 汉王 OCR 版面分析演示

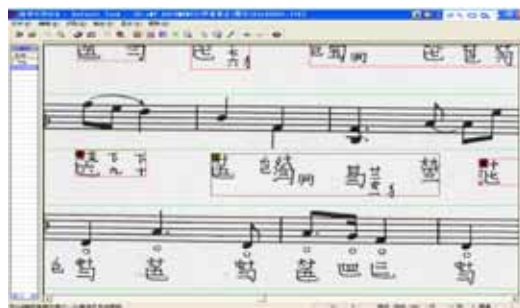


图 10 尚书七号 OCR 版面分析演示

Fig.9 Analysis demonstration of Hanwang OCR layout

Fig.10 Analysis demonstration of Hanwang OCR layout

采用 MiniOCR 进行实验的结果见图 11 与图 12。结果发现当读入图像为较大范围的古琴谱时，可以看出其分析结果是十分混乱的。完全把五线谱与减字谱连在一切，行切分实际上成为块切分。当读入较

小范围的局部琴谱时,也没有正确的切分出两个连接较紧的减字谱,而直接将其忽略过去。同样不能进行图谱的行切分与字切分,只是简单的作出方位标记。



图 11 Mini OCR 段落切分演示 1

Fig.11 Demonstration 1 of Minni OCR
paragraph segmentation



图 12 Mini OCR 段落切分演示 2

Fig.12 Demonstration 2 of Mini OCR
paragraph segmentation



图 13 本文程序行切分演示

Fig.13 program demonstration of test line segmentation



图 14 本文程序字切分演示

Fig.14 program demonstration of word segmentation

然而,采用我们开发的程序系统来进行琴谱切分测试结果见图 13 与图 14。无疑可以看到,对于行切分可以顺利的分析出行投影图,将带有五线谱和减字谱的局部混合琴谱分别切分为两张单纯的五线谱和简字谱图像。字切分也同样,可以顺利的分析出字投影图,并从单行简字中分割出单个简字。

7 结 论

本文给出了一种古琴减字谱图像的文本切分与提取方法及其实现,通过实践验证,很好的实现了琴谱中谱行与字行的分离以及字与字之间的分离。特别是与经典的 OCR 软件相对比,本文所作出的研究成果打破了传统 OCR 软件在处理特殊文本情况下无能为力的盲区,实现了对古琴谱文本切分的高度针对性和有效性,即填补了针对特殊减字谱文字的切分方法空白,同时又对普通的文本切分也具有适用价值。

目前,该切分系统的主要局限方面是只能处理基于五线谱和减字谱较为规范横向排列的琴谱,而对于一些排版较为随意特殊的传统琴谱,切分效果还不理想,需要手动修改。因此还有待于作进一步的完善。

古琴减字谱的自动识别是一个很有意义的全新研究领域,因此这也将是一项长期和细致的工作,具

有较高的难度。我们希望,通过我们的研究,能够推动这方面的研究起到一定的促进作用,从而能够为保护与传播中国古琴文化作出力所能及的贡献。

参考文献:

- [1]周昌乐.手写汉字的机器识别[M].北京:科学出版社,1997.
- [2]陈纯.计算机图像处理技术与算法[M].北京:清华大学出版社,2003.
- [3]孙兆林.MATLAB 6.X 图像处理[M].北京:清华大学出版社,2002.
- [4]周昌乐.一种手写汉字拓扑图表示及获取[J].计算机科学,1996,10.
- [5]周昌乐,赵巍.一种手写汉字文本切分的协同计算方法及其实现[J].河北省科学院学报,1996(增刊).
- [6]Suen C Y, Berthod M, Mori S. Automatic recognition of handprinted characters—the state of the art[J].Proc IEEE,1980,68:469-487.
- [7]Morton J. The interaction of information in word recognition[J]. Psychological Review, 1969,76.
- [8]赵明.手写印刷汉字部件抽取[J].中文信息学报,1988,2(4).

作者简介:

周昌乐,男,厦门大学智能科学与技术系教授,博士生导师,厦门大学艺术认知与计算实验室主任([http:// mac.xmu.edu.cn/](http://mac.xmu.edu.cn/)),主要从事计算琴学、机器心智、认知禅学等方面的研究工作。Email: dozero@xmu.edu.cn